

PENSKORAN POLITOMI DALAM TEORI RESPON BUTIR MENGUNAKAN GRADED RESPONSE MODEL (GRM)

Azhar Rezky Wahyudi*, Anisa, Nasrah Sirajang

Jurusan Matematika, FMIPA, Universitas Hasanuddin, Makassar 90245

*Email: azhar.skripsi@gmail.com

Abstrak

Dalam dunia pendidikan, untuk mengukur kemampuan seseorang dapat menggunakan alat ukur yang biasanya dalam bentuk tes. Dalam pendidikan matematika, terdapat bentuk tes yang umum digunakan termasuk tes pilihan ganda dan tes model uraian. Model dikotomi setiap butir dibagi menjadi dua kategori masalah yakni benar bernilai satu dan salah bernilai nol. Dalam skripsi ini menggunakan model politomi dimana permasalahan dibagi menjadi beberapa kategori (lebih dari dua kategori). Penelitian ini menggunakan desain kelompok paralel ekuivalen. Teori yang digunakan dalam kegiatan penelitian ini adalah teori respon butir (Item response Theory/IRT). Graded Response Model adalah salah satu model dalam IRT untuk memperkirakan kemampuan peserta tes dengan mempertimbangkan parameter peserta, tingkat kesulitan dan isi butir yang membedakan setiap masalah. Dalam skripsi ini, membahas kemampuan peserta tes pada ujian Mid semester mata kuliah matematika dasar yang diberikan kepada mahasiswa Fakultas Perikanan Universitas Hasanuddin tahun ajaran 2011/2012 dengan menggunakan Graded Response Model (GRM).

Kata Kunci: Item Respon Teori (IRT), Graded Response Model (GRM)

PENDAHULUAN

Upaya peningkatan kualitas pendidikan dapat ditempuh melalui peningkatan kualitas pembelajaran dan kualitas sistem penilaiannya. Keduanya saling terkait, sistem pembelajaran yang baik akan menghasilkan kualitas belajar yang baik. Kualitas pembelajaran ini dapat dilihat dari hasil penilaiannya. Selanjutnya sistem penilaian yang baik akan mendorong pendidik untuk menentukan strategi mengajar yang baik dan memotivasi peserta didik untuk belajar lebih baik (Mardapi. 2008;5)

Ada berbagai bentuk tes yang digunakan dalam pendidikan ataupun psikologi. Begitu juga bentuk tes dalam matematika yang misalnya berbentuk objektif dan uraian. Secara umum, pengukuran prestasi belajar pada mata kuliah matematika lebih sering menggunakan tes uraian dibandingkan tes pilihan ganda. Terdapat dua model ujian yang digunakan untuk mengukur keberhasilan peserta tes yaitu; teori ujian klasik dan teori respon butir. Teori ujian klasik bertujuan untuk melihat bagaimana respon dari peserta tes terhadap perangkat tes, sedangkan teori respon butir bertujuan untuk melihat bagaimana respon dari peserta tes terhadap tiap butir soal yang ada dalam perangkat tes. Pada dasarnya, teori respon butir ingin memperbaiki kelemahan yang terdapat pada teori klasik: yakni ketergantungan ukuran ciri butir kepada kelompok peserta dan ketergantungan ukuran ciri peserta kepada kelompok butir tes (Naga, 1992 : 160)

Teori respon butir dikembangkan atas tiga dimensi asumsi dasar. Menurut Hambleton, Swaminathan, & Rogers (1991: 9) dalam (Kana Hidayati, 2002) mengemukakan bahwa asumsi-asumsi yang melandasi teori respon butir yakni:

- a. Asumsi pertama adalah peluang menjawab benar suatu butir tidak dipengaruhi oleh peluang menjawab benar butir yang lain (independensi lokal).
- b. Asumsi kedua, tes mengukur satu dimensi kemampuan (unidimensi).
- c. Asumsi ketiga, pola respon setiap butir tes dapat digambarkan dalam bentuk kurva karakteristik butir.

Hal yang terpenting di dalam teori respon butir adalah penentuan model respon atau karakteristik butir. Model respon harus memenuhi berbagai persyaratan seperti unidimensi, independensi lokal, dan invariansi parameter (Naga, 1992: 175). Teori respon butir membangun suatu model yang menghubungkan karakteristik butir dengan karakteristik peserta. Dengan sejumlah syarat tertentu, model hubungan ini dibuat agar berlaku secara bebas bagi kelompok butir dan kelompok peserta mana saja yang memenuhi syarat itu. Karakteristik butir dan karakteristik peserta dihubungkan oleh model yang berbentuk fungsi atau lengkungan grafik. Sejumlah syarat yang dimaksud dinyatakan dengan sejumlah parameter. Ada parameter butir dan ada pula parameter peserta (Susongko, 2009).

Dalam penelitian ini menggunakan 2 parameter, yakni parameter tingkat kesukaran butir (b) dan parameter daya pembeda (a). Analisis tingkat kesukaran dimaksudkan untuk mengetahui apakah soal tersebut tergolong mudah atau sukar. Tingkat kesukaran adalah bilangan yang menunjukkan sukar atau mudahnya sesuatu soal (Arikunto, 1999:2007). Menurut Allen & Yen (1979: 121), pembagian besarnya indeks kesukaran adalah sebagai berikut:

1. Soal dengan indeks kesukaran 0,00 sampai 0,30 adalah soal sukar
2. Soal dengan indeks kesukaran 0,31 sampai 0,70 adalah soal sedang
3. Soal dengan indeks kesukaran 0,71 sampai 1,00 adalah soal mudah

Parameter lain yang ikut dianalisis dalam penelitian ini adalah parameter daya pembeda. Arikunto (1999 : 211) mengemukakan bahwa daya pembeda soal adalah kemampuan suatu soal untuk membedakan antara siswa yang berkemampuan tinggi dengan siswa yang berkemampuan rendah. Seperti halnya indeks kesukaran soal, maka indeks daya pembeda soal besarnya berkisar antara 0,00 sampai dengan 1,00 (Sudijono, 2008).

Seiring dengan perkembangan metode penskoran terhadap suatu perangkat tes, sekarang teori respon butir telah bisa digunakan dalam penskoran polikotomi (politomi). Dalam penskoran politomi terdapat lebih dari dua nilai yang membentang dari nilai yang terendah sampai nilai yang tertinggi, misalnya nilai yang diberikan pada satu butir tes membentang antara nilai 0 sampai dengan 2 atau nilai yang membentang dengan skala yang lebih luas.

Ada beberapa model penskoran berbentuk politomi. Dari beberapa metode yang ada, salah satu metode yang sering digunakan adalah model GRM (*Graded Response Model*). GRM adalah sistem penskoran dimana tingkat kesukaran tiap kategori pada butir tes disusun secara berurutan sehingga jawaban peserta tes haruslah terurut dari kategori yang rendah hingga kategori yang tinggi.

Terdapat dua perangkat tes yang diuji dalam penelitian ini dengan tujuan untuk mengetahui apakah kedua perangkat tes itu setara atau tidak setara berdasarkan uji homogenitas dan untuk menentukan soal yang layak berdasarkan parameter daya beda dan tingkat kesukaran butir soal dari dua perangkat tes yang digunakan dengan *Graded Response Model* (GRM).

GRM sangat tepat digunakan untuk butir yang memiliki respons kategorikal seperti skala likert. Model respons bergradasi tidak menghendaki kesamaan jumlah kategori respons antar butir. Model respon bergradasi merupakan perluasan dari model 2-PL dimana setiap kategori respon pada butir diperlakukan layaknya butir dikotomi sehingga kurva probabilitas sebanyak jumlah kategori respons.

Dalam model ini terlebih dahulu kita harus mendapatkan fungsi karakteristik operasi (*Operating Characteristic Functions*/OCF) untuk bahan dasar membuat fungsi respon kategori (*Categori Response Functions*/CRF).

OCF dapat dibentuk dengan rumus :

$$P_{ij}^* = \frac{\exp(\alpha_i(\theta_n - \beta_{ij}))}{1 + \exp(\alpha_i(\theta_n - \beta_{ij}))}$$

Dimana : α_i = daya beda pada butir ke-i
 β_{ij} = tingkat kesukaran kategori-i pada butir ke-j
 θ_n = parameter peserta

Persamaan di atas seperti halnya persamaan 2-PL namun masih lebih spesifik lagi, yaitu dalam butir-i terdapat sejumlah j kategori yang masing-masing diestimasi secara terpisah.

OCF tidak dapat dipakai untuk melihat perbandingan probabilitas tiap kategori butir, oleh karena itu maka perlu meneruskan langkah berikutnya dengan menghitung CRF butir.

CRF untuk setiap kategori dinyatakan dengan persamaan sebagai berikut:

$$P_{ij}(\theta) = P_{ij}^*(\theta) - P_{i(j+1)}^*(\theta)$$

Dengan ketentuan bahwa $P_{i0}^*(\theta) = 1$ dan $P_{i(j+1)}^*(\theta) = 0$

Keterangan :

$P_{ij}(\theta)$ = Probabilitas butir – i kategori – j

$P_{ij}^*(\theta)$ = Probabilitas butir – i kategori yang lebih awal

$P_{i(j+1)}^*(\theta)$ = Probabilitas butir – i kategori paling akhir

Perhitungan CRF berdasarkan persamaan di atas dapat kita jabarkan menjadi 3 kategori, yakni sebagai berikut:

Kategori 0 : $P_{10}(\theta) = 1 - P_{11}^*(\theta) = 0 - \frac{\exp(\alpha_1(\theta_n - \beta_{11}))}{1 + \exp(\alpha_1(\theta_n - \beta_{11}))}$

Kategori 1 : $P_{11}(\theta) = P_{11}^*(\theta) - P_{12}^*(\theta) = 0 - \frac{\exp(\alpha_1(\theta_n - \beta_{11}))}{1 + \exp(\alpha_1(\theta_n - \beta_{11}))} - \frac{\exp(\alpha_1(\theta_n - \beta_{12}))}{1 + \exp(\alpha_1(\theta_n - \beta_{12}))}$

Kategori 2 : $P_{12}(\theta) = P_{12}^*(\theta) - 0 = \frac{\exp(\alpha_1(\theta_n - \beta_{12}))}{1 + \exp(\alpha_1(\theta_n - \beta_{12}))} - 0$

METODOLOGI PENELITIAN

Data yang akan digunakan adalah data primer yang berupa nilai yang diperoleh berdasarkan hasil ujian Mid Semester Mata Kuliah Matematika Dasar yang akan diberikan kepada mahasiswa Fakultas Perikanan Universitas Hasanuddin tahun ajaran 2011/2012 dengan variabel penelitian yang digunakan adalah sebagai berikut:

X = perangkat tes pertama

Y = perangkat tes kedua

Metode yang digunakan dalam mengumpulkan dan menganalisis data, adalah sebagai berikut:

1. Penyusunan perangkat tes meliputi:
 - a. Materi yang akan diteskan
 - b. Menentukan alokasi waktu yang digunakan untuk menyelesaikan tes
 - c. Merancang bentuk tes dalam bentuk uraian
 - d. Menentukan jumlah soal tes sebanyak 10 butir soal mid matakuliah kalkulus dasar
 - e. Menyusun perangkat tes dan kunci jawaban yang berguna untuk mempermudah pengoreksian terhadap hasil jumlah responden
2. Langkah-langkah untuk menganalisis data, yang berkaitan dengan tujuan penulisan skripsi ini adalah sebagai berikut:
 - a. Metode dokumentasi untuk mengetahui nama, nomor induk mahasiswa dan jumlah mahasiswa sebagai peserta tes pada tiap kelasnya.
 - b. Mengumpulkan data melalui hasil ujian mid semester fakultas Perikanan Universitas Hasanuddin tahun ajaran 2011/2012.
 - c. Melakukan Penskoran data yang berasal dari skor yang diperoleh peserta tes dengan menggunakan model penskoran GRM (Graded Respons Model)

Penskoran dilakukan dengan model GRM 3 kategori, yakni:

No	Aspek Penilaian	Skor
1	Mahasiswa tidak dapat menyelesaikan langkah pertama dengan benar	0
2	Mahasiswa hanya dapat menyelesaikan langkah pertama dengan benar tetapi tidak dapat menyelesaikan langkah kedua	1
3	Mahasiswa dapat menyelesaikan seluruh langkah dengan benar	2

Tabel 1. Penskoran model GRM

- d. Hasil Penskoran pada model GRM (*Graded Respons Model*) pada tahap di atas kemudian dilakukan pengolahan data untuk memperoleh parameter daya beda (a) dan tingkat kesukaran (b).
- e. Mengevaluasi butir soal berdasarkan data yang telah diperoleh dari tahap sebelumnya

HASIL DAN PEMBAHASAN

Bentuk rancangan penyetaraan tes yang digunakan dalam penelitian ini adalah rancangan kelompok ekuivalen (equivalent-group design). Rancangan kelompok ekuivalen adalah rancangan dari dua perangkat tes yang diberikan pada dua kelompok yang sama yang kemampuannya diasumsikan ekuivalen atau sama.

Dalam penelitian ini, ujian dilakukan sebanyak dua kali. Peserta ujian dibagi menjadi dua kelompok secara acak yang terdiri dari kelompok pertama (K_1) mengerjakan soal X dan kelompok kedua (K_2) mengerjakan soal Y (untuk ujian pertama), kemudian untuk ujian kedua kelompok K_2 mengerjakan perangkat tes Y dan kelompok K_1 mengerjakan perangkat tes X.

Jumlah peserta ujian sebanyak 52 orang, dimana Kelompok pertama (K_1) yang mengerjakan soal X dan kelompok kedua (K_2) yang mengerjakan soal Y terdiri atas 26 mahasiswa.

Jumlah soal pada masing-masing perangkat tes adalah 13 butir soal yang berbentuk soal uraian. Berdasarkan rancangan yang digunakan dalam penelitian ini, masing-masing perangkat tes terdiri dari 13 butir soal uraian dengan skor terendah adalah 0 dan skor tertinggi adalah 10.

Uji Levene

Uji Levene digunakan untuk mengetahui kehomogenan variansi terhadap parameter-parameter yang diuji yakni tingkat kesukaran butir soal dan daya beda soal. Uji Levene dilakukan terhadap setiap pasangan parameter pada setiap perangkat tes X dan Y.

Adapun hipotesis dalam uji Levene yaitu sebagai berikut:

H_0 : tidak ada perbedaan signifikansi variansi antara perangkat tes X dan Y

H_1 : ada perbedaan signifikansi variansi antara perangkat tes X dan Y

Atau

$H_0 : \sigma_1 = \sigma_2$

$H_1 : \sigma_1 \neq \sigma_2$

Taraf nyata (α) yang digunakan pada penelitian ini adalah sebesar $\alpha = 5\%$ (0.05) dan kriteria pengujiannya yaitu H_0 diterima (H_1 ditolak) apabila nilai signifikan uji Levene $> \alpha = 0.05$ dan H_1 diterima (H_0 ditolak) apabila nilai signifikan uji Levene $< \alpha = 0.05$.

Output dari uji Levene menggunakan perangkat lunak SPSS versi 18 yang secara ringkas disajikan pada tabel berikut:

Tabel 2. Nilai signifikansi uji Levene

Ujian	Perangkat Tes		Parameter		Keterangan
			P	A	
Ujian I	X	Y	0,819	0,063	Terima H_0
Ujian II	Y	X	0,087	0,071	Terima H_0

Berdasarkan teori uji levene yang menyatakan bahwa untuk mengetahui sebuah data homogenitas atau tidak maka kita dapat melihat nilai levene statistik dan juga nilai signifikansinya, apabila nilai signifikansinya lebih besar dari $\alpha = 0,05$ maka data tersebut dapat dikatakan homogen.

Dari hasil analisis data maka dapat dilihat bahwa nilai signifikansi untuk tingkat kesukaran butir (P) pada ujian I yaitu sebesar 0.819 (> 0.05) homogen dengan tingkat kesukaran butir (P) pada ujian II dimana nilai signifikansinya sebesar 0.087 (> 0.05). Sedangkan untuk nilai daya beda (A) pada ujian I nilai signifikansinya sebesar 0,063 (> 0.05) homogen dengan nilai daya beda (A) pada ujian II dimana nilai signifikansinya sebesar 0,071 (> 0.05).

Melihat dari hasil uji levene kedua parameter yang dipertimbangkan yaitu parameter daya beda dan tingkat kesukaran butir pada perangkat tes X dan Y pada ujian I dan ujian II maka dapat dikatakan bahwa kedua perangkat tes ini sudah setara dalam mengukur kemampuan peserta ujian. Hal ini dapat dibuktikan dengan melihat hasil uji levene pada kedua parameter, dimana hasil yang didapatkan yaitu bahwa parameter tingkat kesukaran butir (P) pada ujian I homogen dengan tingkat kesukaran butir (P) pada ujian II dan nilai daya beda (α) pada ujian I homogen dengan nilai daya beda (α) pada ujian II.

Estimasi Parameter Butir

Untuk lebih mempermudah analisis maka data skor mentah yang telah diperoleh dari penelitian ini dikonversi menjadi data skor politomi yaitu 0, 1, dan 2. Peserta yang memperoleh skor 0-4 akan dikonversi ke 0, peserta yang memperoleh skor 5-7 akan dikonversi ke 1, dan peserta yang memperoleh skor 8-10 akan dikonversi ke skor 2.

Menentukan Jumlah Benar (Skor) Kelompok Atas (A) Dan Jumlah Benar (Skor) Kelompok Bawah (B)

Seperti yang telah dikemukakan di atas bahwa perangkat tes terbagi menjadi dua, yaitu perangkat tes X dan perangkat tes Y. Pada gelombang I. Perangkat tes X dikerjakan oleh kelompok 1 (K_1X) dan perangkat tes Y dikerjakan oleh kelompok 2 (K_2Y). Setelah itu, kemudian perangkat tes dipertukarkan pada gelombang II, dimana perangkat tes Y dikerjakan oleh kelompok 1 (K_1Y) dan perangkat tes X dikerjakan oleh kelompok 2 (K_2X). Pada masing-masing kelompok kemudian ditentukan jumlah peserta yang menjawab benar pada kelompok atas (A) dan jumlah peserta yang menjawab benar pada kelompok bawah (B). Untuk menentukan peserta kelompok atas (A) dan peserta kelompok bawah (B) diperoleh dengan cara mengurutkan skor peserta ujian dari yang paling tinggi sampai yang paling rendah dibagi menjadi 2 kelompok yakni kelompok teratas dan kelompok terbawah dari jumlah peserta ujian untuk masing-masing kelompok (K_1) dan kelompok (K_2) dengan perbandingan 50 : 50.

Setelah menemukan nilai jumlah benar (skor) kelompok atas (A) dan jumlah benar (skor) kelompok bawah (B) kemudian kita akan melakukan analisis butir soal untuk menemukan nilai tingkat kesukaran butir dan nilai daya beda.

Menghitung Nilai Tingkat Kesukaran

Pada bentuk butir politomi, nilai tingkat kesukaran butir adalah jumlah dari tingkat kesukaran tahap tiap butir. Nilai tingkat kesukaran butir soal pada Kelompok K_1 yang mengerjakan perangkat tes X dan kelompok K_2 yang

mengerjakan perangkat tes Y dapat dilihat pada tabel (3) dan tabel (4), serta kelompok K_1 yang mengerjakan perangkat tes Y dan kelompok K_2 yang mengerjakan perangkat tes X pada tabel (5) dan tabel (6), menggunakan rumus:

$$\delta_{ij} = \frac{n_{ij}}{N_i}$$

Keterangan :

δ_{ij} = tingkat kesukaran tahap j pada butir i

n_{ij} = jumlah skor yang diperoleh peserta yang menjawab pada butir i kategori j

N_i = skor total pada butir i

Tabel 3. Tingkat kesukaran butir soal pada K_1X

Butir	Indeks Kesukaran Kategori			Total	Keterangan
	Kat.1	Kat.2	Kat.3		
1	0	0	0.42308	0.42308	Soal tergolong sedang
2	0	0.01923	0.07692	0.09615	Soal tergolong sukar
3	0	0	0.07692	0.07692	Soal tergolong sukar
4	0	0.11538	0	0.11538	Soal tergolong sukar
5	0	0	0.07692	0.07692	Soal tergolong sukar
6	0	0.01923	0	0.01923	Soal tergolong sukar
7	0	0	0.42308	0.42308	Soal tergolong sedang
8	0	0.01923	0.19231	0.21154	Soal tergolong sukar
9	0	0.01923	0	0.01923	Soal tergolong sukar
10	0	0.09615	0	0.09615	Soal tergolong sukar
11	0	0.11538	0.15385	0.26923	Soal tergolong sukar
12	0	0	0.03846	0.03846	Soal tergolong sukar
13	0	0	0	0	Soal tergolong sukar

Tabel 4. Tingkat kesukaran butir soal pada K_2Y

Butir	Indeks Kesukaran Kategori			Total	Keterangan
	Kat.1	Kat.2	Kat.3		
1	0	0.05769	0.11538	0.17308	Soal tergolong sukar
2	0	0.13462	0.26923	0.40385	Soal tergolong sedang
3	0	0.03846	0.15385	0.19231	Soal tergolong sukar
4	0	0.17308	0.07692	0.25	Soal tergolong sukar
5	0	0.09615	0.11538	0.21154	Soal tergolong sukar
6	0	0.21154	0.03846	0.25	Soal tergolong sukar
7	0	0	0.65385	0.65385	Soal tergolong sedang
8	0	0.01923	0.30769	0.32692	Soal tergolong sedang
9	0	0.05769	0.07692	0.13462	Soal tergolong sukar
10	0	0.03846	0.03846	0.07692	Soal tergolong sukar
11	0	0.11538	0.15385	0.26923	Soal tergolong sukar

Butir	Indeks Kesukaran Kategori			Total	Keterangan
	Kat.1	Kat.2	Kat.3		
12	0	0.09615	0.38462	0.48077	Soal tergolong sedang
13	0	0.01923	0	0.01923	Soal tergolong sukar

Tabel 5. Tingkat kesukaran butir soal pada K_1Y

Butir	Indeks Kesukaran Kategori			Total	Keterangan
	Kat.1	Kat.2	Kat.3		
1	0	0.03846	0.65385	0.692308	Soal tergolong sedang
2	0	0.11538	0.11538	0.230769	Soal tergolong sukar
3	0	0	0.11538	0.115385	Soal tergolong sukar
4	0	0.01923	0.03846	0.057692	Soal tergolong sukar
5	0	0.07692	0.11538	0.192308	Soal tergolong sukar
6	0	0.07692	0.34615	0.423077	Soal tergolong sedang
7	0	0	0.57692	0.576923	Soal tergolong sedang
8	0	0.01923	0.30769	0.326923	Soal tergolong sedang
9	0	0.03846	0.03846	0.076923	Soal tergolong sukar
10	0	0.13462	0.03846	0.173077	Soal tergolong sukar
11	0	0.09615	0.07692	0.173077	Soal tergolong sukar
12	0	0	0.15385	0.153846	Soal tergolong sukar
13	0	0.07692	0	0.076923	Soal tergolong sukar

Tabel 6. Tingkat kesukaran butir soal pada K_2X

Butir	Indeks Kesukaran Kategori			Total	Keterangan
	Kat.1	Kat.2	Kat.3		
1	0	0	0.03846	0.03846	Soal tergolong sulit
2	0	0.01923	0.03846	0.05769	Soal tergolong sulit
3	0	0.05769	0	0.05769	Soal tergolong sulit
4	0	0	0	0	Soal tergolong sulit
5	0	0	0.19231	0.19231	Soal tergolong sulit
6	0	0.01923	0.11538	0.13462	Soal tergolong sulit
7	0	0.01923	0.46154	0.48077	Soal tergolong sedang
8	0	0.05769	0.11538	0.17308	Soal tergolong sulit
9	0	0.03846	0.11538	0.15385	Soal tergolong sulit
10	0	0.03846	0.03846	0.07692	Soal tergolong sulit
11	0	0.17308	0	0.17308	Soal tergolong sulit
12	0	0.05769	0.07692	0.13462	Soal tergolong sulit
13	0	0	0	0	Soal tergolong sulit

Menghitung Nilai Daya Beda

Nilai daya beda pada kelompok K_1 yang mengerjakan perangkat tes X dan kelompok K_2 yang mengerjakan perangkat tes Y pada tabel (3) dan tabel (4), serta Kelompok K_1 yang mengerjakan perangkat tes Y dan kelompok K_1 yang mengerjakan perangkat tes X pada tabel (5) dan tabel (6), ditemukan menggunakan rumus:

$$\alpha = \frac{RA-RB}{N}$$

Keterangan : α = Daya Beda
RA = Rerata Kelas Atas

$$RA = \frac{na}{Na}$$

na = Total skor peserta kelas atas

Na = Jumlah peserta kelas atas

RB = Rerata Kelas Bawah

$$RB = \frac{nb}{Nb}$$

nb = Total skor peserta kelas bawah

Nb = Jumlah peserta kelas bawah

N_{\max} = Skor Maksimum tiap butir

Berikut akan disajikan tabel nilai daya beda untuk K_1X , K_2Y , K_1Y , dan K_2X

Tabel 7. Nilai daya beda K_1X

Butir	Daya Beda	Penjelasan Daya Beda Soal
1	0,38	Soal diterima dan diperbaiki
2	0,19	Soal ditolak atau dibuang
3	0,15	soal ditolak atau dibuang
4	0,08	soal ditolak atau dibuang
5	0,15	Soal ditolak atau dibuang
6	0,04	soal ditolak atau dibuang
7	0,52	soal tergolong baik dan boleh digunakan
8	0,27	Soal diperbaiki
9	0,04	soal ditolak atau dibuang
10	0,12	soal ditolak atau dibuang
11	0,46	soal tergolong baik dan boleh digunakan
12	0,08	soal ditolak atau dibuang
13	0	soal ditolak atau dibuang

Tabel 8. Nilai daya beda K_2Y

Butir	Daya Beda	Penjelasan Daya Beda Soal
1	0,19	Soal ditolak
2	0,42	Soal baik dan dapat digunakan
3	0,31	Soal diterima dan diperbaiki
4	0,04	soal ditolak

Butir	Daya Beda	Penjelasan Daya Beda Soal
5	0,12	Soal ditolak
6	0,27	Soal diperbaiki
7	0,23	soal diperbaiki
8	0,65	Soal baik dan dapat digunakan
9	0,27	Soal diperbaiki
10	0,15	Soal ditolak
11	0,38	Soal diterima dan diperbaiki
12	0,5	Soal baik dan dapat digunakan
13	0,04	Soal ditolak

Tabel 9. Nilai daya beda K₁Y

Butir	Daya Beda	Penjelasan Daya Beda Soal
1	0,38	Soal diterima dan diperbaiki
2	0,31	Soal diterima dan diperbaiki
3	0,23	Soal diperbaiki
4	0,12	Soal ditolak
5	0,23	Soal diperbaiki
6	0,31	Soal diperbaiki
7	0,54	Soal diterima dan diperbaiki
8	0,5	Soal diterima dan diperbaiki
9	0,15	Soal ditolak
10	0,19	Soal ditolak
11	0,35	Soal diterima dan diperbaiki
12	0,31	Soal diterima dan diperbaiki
13	0,15	Soal ditolak

Tabel 10. Nilai daya beda K₂X

Butir	Daya Beda	Penjelasan daya Beda Soal
1	0,08	Soal ditolak
2	0,12	Soal ditolak
3	0,12	Soal ditolak
4	0	Soal ditolak
5	0,38	Soal diterima dan diperbaiki
6	0,27	Soal diperbaiki
7	0,42	Soal baik dan dapat digunakan
8	0,27	Soal diperbaiki
9	0,08	Soal ditolak
10	0,15	Soal ditolak
11	0,27	Soal diperbaiki
12	0,27	Soal diperbaiki

Butir	Daya Beda	Penjelasan daya Beda Soal
13	0	Soal ditolak

Parameter Peserta

Jika kita memperhatikan pelaksanaan, maka pengestimasiannya ini dilakukan untuk mengetahui bagaimana ciri peserta atau bagaimana kemampuan peserta pada uji tes yang telah dilakukan. Dalam uji tes, ciri peserta adalah tetap sekalipun butir uji tes diganti. Parameter peserta yang mengikuti ujian tes ini berbeda-beda berdasarkan tingkat kemampuannya.

Tabel 11. Parameter peserta

No	Nama Peserta	NIM	Parameter Peserta
1	Peserta 1	L23109256	0.5
2	Peserta 2	L23111006	0.5
3	Peserta 3	L23111252	0.9
4	Peserta 4	L23111253	0.7
5	Peserta 5	L23111264	0.8
6	Peserta 6	L23111257	0.8
7	Peserta 7	L23111003	0.8
8	Peserta 8	L23111021	0.8
9	Peserta 9	L23111015	0.7
10	Peserta 10	L23111255	0.5
11	Peserta 11	L23111017	0.5
12	Peserta 12	L23111259	0.8
13	Peserta 13	L23111018	0.9
14	Peserta 14	L23111019	0.3
15	Peserta 15	L23111266	0.5
16	Peserta 16	L23111002	0.7
17	Peserta 17	L23111261	0.7
18	Peserta 18	L23111270	0.8
19	Peserta 19	L23111009	0.5
20	Peserta 20	L23111014	0.6
21	Peserta 21	L23111008	0.7
22	Peserta 22	L23111013	0.4
23	Peserta 23	L23111101	0.5
24	Peserta 24	L23111001	0.9
25	Peserta 25	L23111269	0.7
26	Peserta 26	L23111020	0.5
27	Peserta 27	L23110274	0.6
28	Peserta 28	L23109260	0.4
29	Peserta 29	L23111010	0.5
30	Peserta 30	L23111004	0.7
31	Peserta 31	L23111256	0.7
32	Peserta 32	L23111102	0.5

33	Peserta 33	L23111016	0.5
34	Peserta 34	L23111258	0.8
35	Peserta 35	L23111254	0.7
36	Peserta 36	L23111011	0.6
37	Peserta 37	L23111251	0.7
38	Peserta 38	L23111901	0.5
39	Peserta 39	L23111268	0.5
40	Peserta 40	L23111022	0.8
41	Peserta 41	L23109274	0.8
42	Peserta 42	L23111260	0.1
43	Peserta 42	L23111601	0.5
44	Peserta 44	L23111272	0.5
45	Peserta 45	L23111267	0.5
46	Peserta 46	L23111007	0.7
47	Peserta 47	L23111012	0.5
48	Peserta 48	L23111273	0.6
49	Peserta 49	L23111265	0.7
50	Peserta 50	L23111005	0.9
51	Peserta 51	L23111902	0.5
52	Peserta 52	L23111262	0.7

Menentukan Nilai *Graded Response Model (GRM)*

Dalam model ini terlebih dahulu kita harus mendapatkan fungsi karakteristik operasi (*Operating Characteristic Functions/OCF*) untuk bahan dasar membuat fungsi respon kategori (*Categori Response Functions/CRF*). OCF dapat dibentuk dengan rumus :

$$P_{ij}^* = \frac{\exp(\alpha_i(\theta_n - \beta_{ij}))}{1 + \exp(\alpha_i(\theta_n - \beta_{ij}))}$$

Dengan memasukkan nilai α_1 (tiap butir) yang telah diperoleh dari table 7, tabel 8, tabel 9, dan table 10, kemudian memasukkan nilai θ (tiap peserta) yang diperoleh dari tabel 11, kemudian memasukkan nilai β_{ij} (tiap butir pada setiap kategori, dimana terdapat 3 kategori yakni kategori 0, kategori 1, dan kategori 2) maka akan diperoleh nilai P_{ij}^* untuk tiap butir pada tiap kategori.

Perhitungan CRF berdasarkan persamaan di atas kita jabarkan menjadi 3 kategori, yakni sebagai berikut:

$$\text{Kategori 0 : } P_{10}(\theta) = 0 - P_{i1}^*(\theta) = 0 - \frac{\exp(\alpha_1(\theta_n - \beta_{11}))}{1 + \exp(\alpha_1(\theta_n - \beta_{11}))}$$

$$\text{Kategori 1 : } P_{11}(\theta) = P_{i1}^*(\theta) - P_{i2}^*(\theta) = \frac{\exp(\alpha_1(\theta_n - \beta_{11}))}{1 + \exp(\alpha_1(\theta_n - \beta_{11}))} - \frac{\exp(\alpha_1(\theta_n - \beta_{12}))}{1 + \exp(\alpha_1(\theta_n - \beta_{12}))}$$

$$\text{Kategori 2 : } P_{12}(\theta) = P_{i3}^*(\theta) - 1 = \frac{\exp(\alpha_1(\theta_n - \beta_{12}))}{1 + \exp(\alpha_1(\theta_n - \beta_{12}))} - 1$$

Hasil pengolahan data dari fungsi karakteristik operasi (*Operating Characteristic Function/OCF*) untuk tes ujian pertama dapat dilihat pada lampiran 2, sedangkan hasil pengolahan data dari fungsi karakteristik operasi (OCF) untuk tes kedua dapat dilihat pada lampiran 3.

Setelah mendapatkan nilai OCF untuk setiap peserta pada setiap butir dalam setiap kategori, selanjutnya kita menghitung fungsi respons kategori (*Category Response Function/CRF*).

CRF untuk setiap kategori dinyatakan dengan rumus sebagai berikut:

$$P_{ij}(\theta) = P_{ij}^*(\theta) - P_{i(j+1)}^*(\theta) = 0$$

Dengan ketentuan bahwa $P_{i0}^*(\theta) = 1$ dan $P_{i(j+1)}^*(\theta) = 0$

Persamaan di atas diberlakukan pada semua kategori dalam butir, dengan menetapkan bahwa P_{ij} terendah adalah 0 dan P_{ij} tertinggi adalah 1. Probabilitas kategori paling awal P_{i0}^* adalah 1 dan kategori paling akhir $P_{(j+1)}^*$ sama dengan 0. Jika sebuah butir dengan 3 kategori maka kita perlu membuat satu kategori bayangan yang nilainya adalah 1.

Berdasarkan hasil pengolahan data, maka dapat ditarik beberapa kesimpulan yakni:

1. Berdasarkan uji homogenitas variansi dari parameter tingkat kesukaran butir (P) dan parameter daya beda (A) kedua perangkat tes maka dapat disimpulkan bahwa perangkat ujian Mid Semester Mahasiswa Fakultas Perikanan Universitas Hasanuddin Tahun ajaran 2011/2012 adalah setara untuk perangkat tes X dan Y,
2. Soal yang baik adalah soal yang memiliki tingkat kesukaran butir yang sedang dan nilai daya beda yang tinggi. Berdasarkan hasil ujian Mid Semester yang telah diperoleh oleh Mahasiswa Perikanan Universitas Hasanuddin maka didapatkan beberapa soal yang baik untuk kedua perangkat tes, dimana untuk perangkat tes X soal yang baik adalah soal nomor 7 dan untuk perangkat tes Y tidak ada soal yang bisa langsung digunakan tetapi sebaiknya diperbaiki terlebih dahulu seperti pada soal 1, 8, dan 12,
3. Perbandingan probabilitas tiap kategori yang sama dalam setiap butir dengan menggunakan model GRM cenderung menghasilkan nilai yang sama besar antara kategori 0 dengan kategori 1 di tiap butirnya dengan parameter kemampuan peserta yang ada.

DAFTAR PUSTAKA

- Allen & Yen. (1979). *Introduction to Measurement Theory*. Belmont, CA : Wadsworth, Inc.
- Arikunto. 2007. **Dasar – dasar Evaluasi Pendidikan**. Jakarta : Bumi Aksara.
- Hidayati, Kana. 2002. <http://staff.uny.ac.id/sites/default/files/penelitian/KanaHidayati,M.Pd/PenerapanIRTdalampenyetaraan.pdf>. (diakses tanggal 22 April 2012).
- Mardapi, Djemari. 2008. **Teknik Penyusunan Instrumen Tes dan Non Tes**. Yogyakarta: Mitra Cendekia
- Naga, Dali S. 1992. **Pengantar Teori Sekor pada Pengukuran Pendidikan**. Jakarta: Gunadarma
- Sudijono, Anas. 2008. **Pengantar Evaluasi Pendidikan**. Jakarta: Grafindo Persada
- Susongko, Purwo. 2009. **Perbandingan Keefektifan Bentuk Tes Uraian dan Teslet dengan Penerapan Model GRM**. Pascasarjana UNY.